

国立国語研究所学術情報リポジトリ

漢語の仮名表記：実態と背景

著者	間淵 洋子
雑誌名	言語資源活用ワークショップ発表論文集
巻	1
ページ	201-213
発行年	2017
URL	http://doi.org/10.15084/00001475

漢語の仮名表記 —実態と背景—

間淵 洋子（明治大学国際日本学研究科・日本学術振興会）[†]

A Corpus Based Study of Sino-Japanese Words written in *Kana*: Actual Condition and Background

MABUCHI Yoko (Meiji University / Society for the Promotion of Science)

要旨

本発表は、本来漢字で表記されるはずの漢語が平仮名や片仮名で表記される事象を取り上げ、『現代日本語書き言葉均衡コーパス』（以下、「BCCWJ」と表記）を用いて、その実態と背景を明らかにすることを目的とする。

BCCWJの網羅的な漢語の表記実態調査に基づき、個々の語の仮名表記率から、仮名表記が、主たる表記である語、ある程度一般的である語を特定した上で、仮名表記の定着度合いに、**字体特徴**（常用漢字表外字・音を含む語は仮名表記率が高いが、表内字でも仮名表記率の高い語がある）、**語の出現状況**（語彙レベルが高い語ほど仮名表記率は低い）、**音声変位形の有無**（「格好」に対する「カッコ」のような音転訛形を持つ語は仮名表記率が高い）、**意味分野**（動植物や食物の分野では仮名表記率が高い）、**品詞**（副詞用法を持つ語は仮名表記率が高い）、**レジスター**（Web媒体は仮名表記率が低い）等との関連性が見られることを示す。また、字体特徴にかかわらず、意味分野や品詞において特定の語彙群に同様の傾向が見られるのは、表記選択に類似性に基づく合理化作用が働くことによると主張する。

1. はじめに

日本語は、漢字、平仮名、片仮名、アルファベットといった多様な文字を持ち、最も複雑な表記体系を持つ言語だと言われる。しかし、その複雑さの中には、和語の主要な意味を担う部分や漢語は漢字で、送り仮名や文法的な役割を表す付属語は平仮名で、外来語を片仮名で、外国語をアルファベットで、といった原則的な役割分担が定まっており、これら多様な文字種は日本語文章の効率的な理解に欠かせないものとなっている。また、それぞれの文字種が持つ感情的な意味を加えたり、あえて原則を外すことで新規性や特殊性を持たせたりといった、表記戦略を取ることで、日本語表現を極めて豊かなものとしている。

このような字種の多様性と語の表記の関連性について言及したこれまでの研究の多くは、文字種と語種の対応関係に目を向け、主に外来語以外で片仮名表記される事象を取り上げ、その要因や効果、機能を述べたものであった（中山 1998, 成田・榊原 2004, 臼木 2008, 柏野 2014 など）。これらの先行研究により、外来語以外で片仮名表記される語には、①常用漢字表外漢字、表外音訓、表外熟字訓を含む語、②動植物名を表す語、③オノマトペ、が多いこと、また、④特殊な語義の書き分け（意味の限定や専門用語的な用法など）、⑤強調、⑥片仮名の持つ感情的意味の付加、といった表記戦略により、片仮名表記が選択されることが明らかにされている。一方で、特段の効果や機能に寄らない表記の慣習によるものとして、その先に踏み込むことなく残される語があり、これらの表記選択のメカニズムが明らかになっているとは言い難い。また、実際にどのような語に仮名表記の選択される

[†] mabuchi@meiji.ac.jp

慣習があるのかという個々の語の実態把握・解明についても、十分に尽くされてはいない。

そこで、本発表では、特に 2 字漢語を例として、コーパスを用いて網羅的に漢語における仮名表記の実態を調査し、その背景について多角的な考察を試みる。

2. 研究方法

2.1 コーパス

本研究では、漢語の仮名表記の実態をできる限り網羅的に捉えるために、国立国語研究所が 2011 年 8 月に公開した『現代日本語書き言葉均衡コーパス』(BCCWJ)を用いた調査を行う。BCCWJ は、日本語研究ばかりでなく、日本語教育、国語教育、辞書編纂、心理学・認知科学、言語政策といった多様な研究分野への応用を目指し、綿密な設計により構築された、日本初となる、また唯一の、大規模バランストコーパスである(前川 2008)。代表性を有する多種多様な書き言葉が、大量に、かつ、全てに形態論情報が付された状態で収録されており、本研究のように探索的にできるだけ多様で大量の語例を収集するためには、まさに好適なデータであると言える。

2.2 調査対象語の抽出

漢語の仮名表記実態把握を試みるにあたって、全ての漢語を漏れなく抽出し分析を行うことが望ましいのは明らかであるが、本研究では、敢えて 2 字漢語に限って分析対象として扱うこととした。その理由は、概ね以下の 4 点である。

- 1) 漢語は語構成によって、日本古来の語彙(和語)との熟合度が異なっており、特に、1 字漢語の和語との熟合度は、2 字以上の漢語のそれとは大きな隔りがある。熟合度の違いは文法的な振る舞いや、表記にも大きく影響を及ぼしており、1 字漢語と 2 字以上の漢語とは、それぞれ別途分析の必要がある。
- 2) 上記熟合度の高さなども関連して 1 字漢語は、「する」などの和語と、時には連濁を伴い分ちがたく結合することがある。1 字漢語と 2 字漢語は、コーパスでの単位認定においても大きな差異があり、「愛する」「興ずる」「処する」といった 1 語の混種語として扱われる。そのため、他の漢語と同様の方法で網羅的に収集するのが難しい。
- 3) 一方 2 字漢語は、1 字漢語が和語と固く結束するのと同様に、漢字 2 字が離れがたく熟合し意味を持ち、これが語基となって、他の 1 字漢語や 2 字漢語と結びついて新たな語を形成する。漢語において最も基本となる形態であり、語数も多く使用も多い。よって、2 字漢語の実態を把握しておくことは、漢語全体の実態把握の基礎となる。
- 4) 2 字漢語は漢字 2 字が強く結合しており、それ故に語の理解・把握において漢字が担う役割が大きい。漢字そのものが重要な役割を果たしている場合には、本来漢字表記を捨てて仮名表記を選択する動機に乏しいと思われるため、それでもなお仮名表記が選択される背景を調査することは、日本語の表記・語彙体系における漢語の位置づけを探る手掛かりになる。

上記の理由から、本研究では 2 字漢語を調査対象とするが、その抽出には、Web 上のコーパス検索アプリケーション「中納言」を用いた¹。「中納言」の検索においては、以下の条件を設けて、調査対象語として、語彙素が漢字 2 文字からなり、書字形に仮名を含む漢語を網羅的に取り出した。

¹ 現代日本語書き言葉均衡コーパス (非 NumTrans 版) 中納言 2.2.0 を用いた。

キー条件：語種「漢」（漢語），語彙素「[一-ㇿ=][一-ㇿ々=]」，書字形「%[あ-んア-ン]%」
 検索対象：「教科書（OT）」²を除く全てのレジスター

その結果，BCCWJにおいて異なり語数で約3,500語の漢語が抽出された。その中から，数詞や，仮名表記の粗頻度合計が3以下の約1,750語（例：「大学/だいがく」「内容/ないよう」「安全/あんぜん」「選手/せんしゅ」「会議/かいぎ」など）を分析対象外とした。これは，仮名表記の粗頻度が極めて低いものの場合，出現例が特異な例であり，仮名表記選択の背景分析に寄与しない可能性が高いためである。更に，残った半数について各語の総出現頻度（延べ語数）を計測し，それが100に満たない約350語（例：「変挺/へんてこ」「滅法/めっぽう」「鮫鰐/アンコウ」「鷺鳥/ガチョウ」など）も，分析対象外とした。これは，出現頻度が低いものの場合，後節で検討する様々な分類カテゴリやレジスターによる比較を行う際に，計量分析に耐える十分なサンプル数を確保できない可能性があるためである。

また，2音節の2字漢語が平仮名2字で表記されている場合は，解析の誤りであることが多いため，一部表記例を確認し，誤解析が多いものについては，やはり分析対象から外した（「基地/きち」「意気/いき」「未知/ミチ」「磁気/じき」「恣意/しい」など）。以下のような例である（以下，例文においては，注目する語・表記を太字とし，下線を施す）。

- (1) 携帯電話メーカーであゆがイメージキャラクターを勤める。

（解析結果：アユ,阿諛,名詞-普通名詞-サ変可能）³

【出典】BCCWJ サンプルID：PB27_00045 実著者不明・H.A シスターズ(著)『F or A』2002年

これにより，最終的な分析対象は，異なり1,075語，2,835表記，延べ語数4,022,616語となった。

3. 調査結果

3.1 仮名表記率

抽出した1,075語については，出現総頻度に対する仮名表記頻度（平仮名表記と片仮名表記を合わせたもの）の比率である「仮名表記率」を求め，比率により以下に層別した。

特：75%以上 高：50%以上 75%未満 中：25%以上 50%未満 低：25%未満

それぞれの層に含まれる語数と，語例として出現総頻度数の高いものから上位10語を「仮名表記頻度 / 出現総頻度 / 仮名表記率」と共に示すと以下の通りである（表1）。

² 本研究では漢語の仮名表記を扱うが，学習段階の児童・生徒用の教科書においては，未習漢字を仮名で表記する，一般社会における表記とは異なる特殊な事情に基づく仮名表記が含まれているため，調査対象から外すこととした。ただし，調査対象とした書籍や雑誌においても，一部に幼児・児童・生徒向けの作品・記事が含まれているため，左記の特殊な事情に基づく仮名表記は含まれる可能性がある。

³ 全206例中，正しい解析結果は，書字形「阿諛」の23例のみ，仮名表記例「あゆ」183例は全て人名で誤解析であった。

表 1 仮名表記率カテゴリ別語例

カテゴリ	所属語数	所属語例
特 (75%以上)	59	勿論(19,459/20,902/93.1%), 沢山(12,918/15,543/83.1%), 所為(8,272/8,390/98.6%), 御免(3,991/4,339/92.0%), 折角(3,002/3,328/90.2%), 段々(2,861/3,301/86.7%), 林檎(2,102/2,384/88.2%), 無論(1,646/2,187/75.3%), 胡麻(1,767/2,119/83.4%), 大分(1,640/1,913/85.7%)
高 (50%以上)	53	奇麗(8,118/11,677/69.5%), 多分(4,421/7,975/55.4%), 馬鹿(4,761/7,914/60.2%), 是非(5,438/7,909/68.8%), 随分(3,256/4,657/69.9%), 大抵(2,362/3,242/72.9%), 蛋白(2,127/3,106/68.5%), 途端(2,041/3,064/66.6%), 怪我(1,451/2,855/50.8%), 人参(1,445/1,999/72.3%)
中 (25%以上)	81	本当(10,155/35,069/29.0%), 格好(3,151/6,425/49.0%), 大体(2,862/5,985/47.8%), 御飯(1,687/5,560/30.3%), 挨拶(1,407/5,308/26.5%), 普段(1,366/5,208/26.2%), 御覧(1,333/4,283/31.1%), 味噌(1,350/3,843/35.1%), 漫画(1,515/3,523/43.0%), 醤油(1,323/3,260/40.6%)
低 (25%未満)	894	自分(549/110,611/0.5%), 問題(36/70,995/0.1%), 必要(50/70,110/0.1%), 時間(79/67,391/0.1%), 関係(59/57,073/0.1%), 人間(146/42,308/0.3) %, 生活(9/41,372/0.02%), 会社(13/41,115/0.03%), 研究(4/39,936/0.1%), 意味(214/38,966/0.5%)

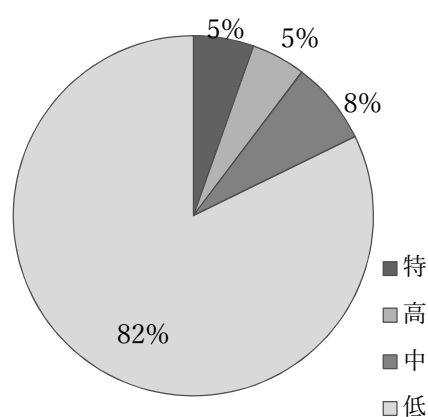


図 1 仮名表記率カテゴリの分布

表 1, 図 1 より, 仮名表記例を持つ語の約 10%は, 仮名表記の比率が 5 割以上あり, 語の優勢な表記が仮名表記であることが分かった。これらは, 仮名表記されることが慣用となっている語とみなすことができる。

また, 8%ほどの語で, 漢字表記が優勢であるものの, ある程度, 仮名表記が一般的に用いられていることが分かった。これらは, 仮名表記を選択している用例と選択していない用例を分析することで, 仮名表記選択の背景についてより細かく観察することができる可能性がある。

そこで, 次に, 仮名表記率は何に起因するか, どのような言語的事象と相関が高いのかを確認するために, いくつかの視点から, 仮名表記率の差異について検討する。

3.2 常用漢字表と仮名表記率の相関

漢語を漢字で表記するか仮名で表記するか, という選択において, 最も関連が深いと思

われるのは、語を構成する漢字やその読みが常用漢字表に含まれているものか否かという点である。常用漢字は、公用文や新聞社・出版社での表記基準となっているため、漢語内に表外字が含まれる場合は、仮名に置き換えて表記されるのが表記原則だからである。

そこで、語を構成する漢字に、常用漢字表⁴に含まれない字（音）が含まれる漢語を「常用外」、常用漢字のみからなる漢語を「常用内」として層別し、各層における仮名表記率を求めたところ、表外字を含む漢語の平均は 20.9%、表内字のみからなる漢語の平均は 3.7%と大きく開きがあった。

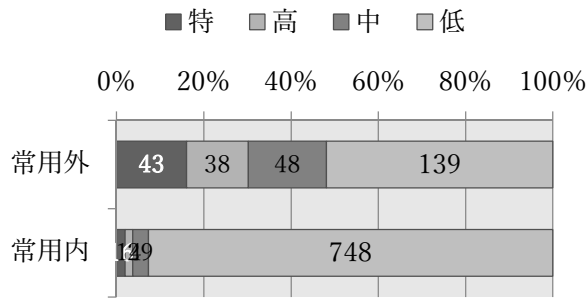


図2 常用漢字表と仮名表記率

また、3.1 節に示した表記率による層別の分布を示した図2からは、表内字のみの漢語は圧倒的に仮名表記率が低い（約 93%）、表外字を含む漢語では、約半数が仮名表記率「中」以上、残りの半数が「低」であり、常用漢字かどうかは仮名表記率に大きく関わるものの、表外字を仮名表記とすることには、かなり大きな幅があることが分かった。

3.3 出現頻度と仮名表記率の相関

前節において、仮名表記率は、常用漢字かどうかによって一意に決まるものではないことを示した。そこで、次に、語の使用状況と仮名表記率との関わりについて見ていきたい。

表2 語彙レベルの基準・範囲と所属語数

レベル	基準値	頻度範囲	累積%	語数	語例（斜体は常用外、太字下線は仮名率「中」以上）
A	60%	9015・	60.01%	131	自分、問題、必要、時間、関係、人間、生活、会社、研究、意味、学校、 <u>本</u> 当、利用、現在
B	80%	1809-9014	80.00%	280	相当、材料、職員、移動、細胞、次第、気分、天皇、携帯、 <u>所</u> 為、興味、以来、距離、義務
C	90%	464-1808	90.01%	291	休憩、衣装、電波、寄付、体操、最悪、景色、 <u>親</u> 戚、親切、頻繁、 <u>稽</u> 古、感心、悪魔、 <u>鯨</u> 鮓
D	95%	157-463	95.00%	271	氾濫、 <u>途</u> 轍、墮落、白鳥、母音、催促、 <u>火</u> 燧、大麻、 <u>褒</u> 美、 <u>傲</u> 慢、変態、 <u>暖</u> 簾、伝言、繁盛
E	99%	20	99.04%	102	漆喰、扶持、面相、躁鬱、 <u>野</u> 暮、飛脚、銀山、劍幕、幻滅、 <u>睡</u> 蓮、 <u>咀</u> 嚼、 <u>蹂</u> 躪、母艦、 <u>軋</u> 轢

語の使用状況は、その語の日本語語彙における重要度や馴染み度（親密度）と相関がある（天野・近藤 2000、寺田・田中 2008 等）。高頻度の語の多くは、日本語の語彙において中心的・基本的な語であり、それらの語を構成する漢字は常用漢字に収録されている可能

⁴ 調査データである BCCWJ は、2005 年までに出版された新聞・雑誌・書籍と、2008 年までに収集された Web データに基づくため、ここでは、2010 年に改訂された現行の常用漢字ではなく、1981 年内閣告示による常用漢字表を用いた。

性が高く、また表記や用法に大きな揺れはないことが推測される。一方、頻度の低い語は、馴染みが薄く、場合によっては難解・難読の語である可能性が高く、情報伝達の必要性から可読性の高い仮名表記が選択されるという背景が推測される。

そこで、実際に語の使用状況を確認するために、調査対象語を「語彙レベル」により層別し使用状況のラベルを与え、仮名表記率との相関を見る。「語彙レベル」は、語の使用頻度の累積度数による「カバー率」（対象データの延べ語数に対して累積頻度が占める比率）に基づき与えられる（田中 2013）。ここでは、田中 2013 による「語彙レベル」の求め方を参考に、独自に基準を設けてラベルを与え（表 2）、仮名表記率との関連を見ていく。

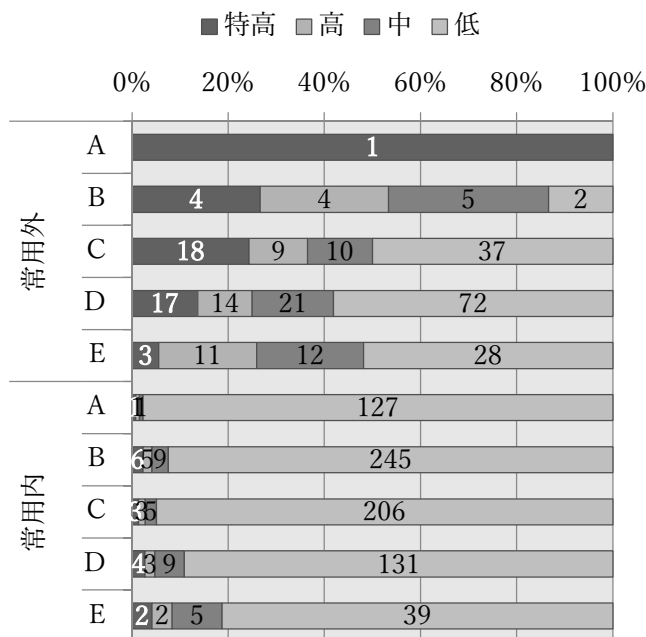


図 3 語彙レベルと仮名表記率の分布

図 2 に、常用漢字かどうかの層別で、語彙レベルごとの仮名表記率の分布を示す。表内字のみからなる漢語においては、総じて仮名表記率が低い中で、語彙レベルが高いほど仮名表記率は低く、語彙レベルが低いほど仮名表記率は高くなる傾向が見られる。

一方、表外字を含む漢語は高レベルの語が少ないものの、語彙レベルが高いほど仮名表記率は高い。この結果は、語の一般性が高いほど、ルールに基づく表記（表内字は漢字表記、表外字は仮名表記）が選択され、一般性が低いほど、それと異なる表記が許容されるという表記傾向の表れと見られる。

3.4 音転訛形と仮名表記率の相関

次に、通常の漢字音からの逸脱により、漢字と語形（音）とに乖離が見られる場合には、仮名表記される頻度が高くなると予想し、音声転訛形を持つか否かによる分析を行った。

(2) 私どもからすると、正直いって、対処法とか考えているだけ面倒くさい。

【出典】 BCCWJ サンプル ID : PB20_00106, 青田吉弘(著)『情報化社会対話集』2002 年

(3) 毎日毎日水をとり替えてめんどくさいでしょ

【出典】 BCCWJ サンプル ID : PB27_00182, 渡辺襄(著)『窯焚き三昧』2002 年

音声転訛形とは、標準形で現れる例(2)に対して、例(3)に見られる語末長音短呼形（面倒：メンドウ→メンド）のようなものや、融合形（絶対：ゼッタイ→ゼッター）などを指すが、加えて、一時的に現れる促音付加（最高：サイコウ→サイッコウ）や、意図的な連母音の長音化（携帯：ケイタイ→ケータイ）も含めた。

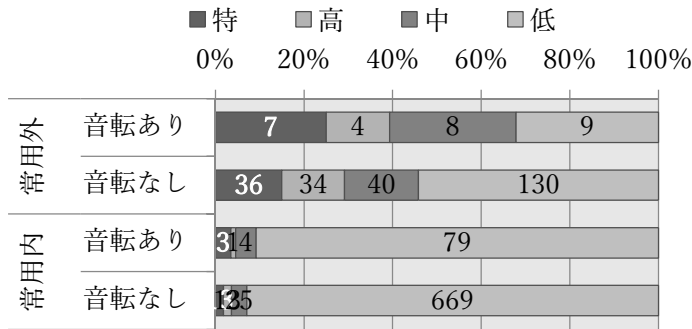


図4 音声転訛形の有無と仮名表記率の分布

図4を見ると、常用漢字表外字を含む漢語では、音声転訛形のある語で、より仮名表記率の高い語が高い傾向にあることが分かる。ただし、常用漢字のみからなる漢語の場合は、そのような差が見られなかった。

3.5 語義分野と仮名表記率の相関

語の意味分野と表記の関連性については、既に先行研究によって動植物名が片仮名表記される傾向にあることが明らかにされているが、それらと常用漢字表との関連性、また、動植物名以外の仮名表記率に関与する分野の有無について検討するために、語義分野と仮名表記率の相関を見た。

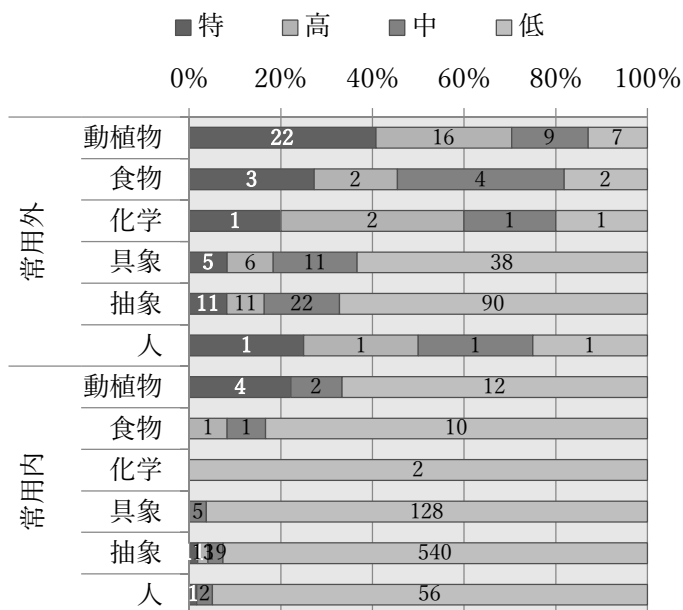


図5 語義分野と仮名表記率の分布

まず、常用漢字のみからなる語に関しては、全体的に仮名表記率が低いものの、他分野に比して動植物名（「昆布」「隠元」など）で仮名表記率が高い「特」「高」の割合が高い。表外字でも同様に、動植物、加えて食物（「味噌」「鰻」「煎餅」など）に関連する語彙で仮名表記率の高い語の割合が多く、抽象概念を表す語（「完璧」「贅沢」「傲慢」「躁鬱」など）で仮名表記率が低い。

通常漢字で表記されるべき、常用漢字のみからなる語で、動植物名に仮名表記率の高い語がある要因は、語義的な分野としての動植物名に表外字や外来語が多いほか、学術用語として用いられる

際に片仮名で表記される用字法が浸透していることの影響とも思われる。また、動植物名は、例(4)で常用漢字のみからなる「大根」が片仮名表記されているように、他の語と列挙される場合、常用漢字か否かに依らず、統一的に仮名表記が用いられることが少なくない。このような表記の統一化・合理化が動植物語彙における仮名表記率の高さに関与しているのではないかと。また同様に、動植物名は食品となるものが多く、意味分野が近接していることによって、動植物ではない食物の語彙についても、上述の統一化・合理化による仮名表記選択がなされている可能性が高い。

(4) ・ダイコン・コマツナ・コマツナ・ネギ・ハクサイ・ハス・ハウレンソウ・芽キャベツ・

ヤマイモ・ユリネ

【出典】 BCCWJ サンプル ID: PB15_00156, 田中元(著)『熟年世代からの元気になる「食生活」の本』 2001 年

- (5) 食料は、スイカ 1 0 0 0 円、花火 1 0 0 0 円、肉 5 1 6 円、たまご・たまねぎ・しょうゆ・さとう・みそ・シーチキン 1 3 0 0 円、きゅうり・とまと・レタス 7 0 0 円、

【出典】 BCCWJ サンプル ID: LBq7_00020, 菅原道彦(著)『あそびの達人』 2002 年

3.6 品詞と仮名表記率の相関

表 3 コーパスの品詞情報と品詞ラベル

コーパスの品詞情報	正規化したラベル
名詞-普通名詞-サ変可能	名詞・動詞
名詞-普通名詞-サ変形状詞可能	名詞・動詞・形状詞
名詞-普通名詞-形状詞可能	名詞・形状詞
名詞-普通名詞-副詞可能	名詞・副詞
その他「名詞」から始まる品詞	名詞
形状詞-タリ／形状詞-一般	形状詞
接尾辞-名詞的-一般	名詞
副詞	副詞

次に、それぞれの漢語に付加されている品詞情報によって漢語を層別し、仮名表記率との相関を見る。コーパスの形態論情報として付与された品詞については、表 3 に示す正規化を行い、「結構」のように複数品詞を持つ語（形状詞-一般、副詞、名詞-普通名詞-一般）は同様に「名詞・副詞・形状詞」のように品詞を繋いだラベルを付けた。

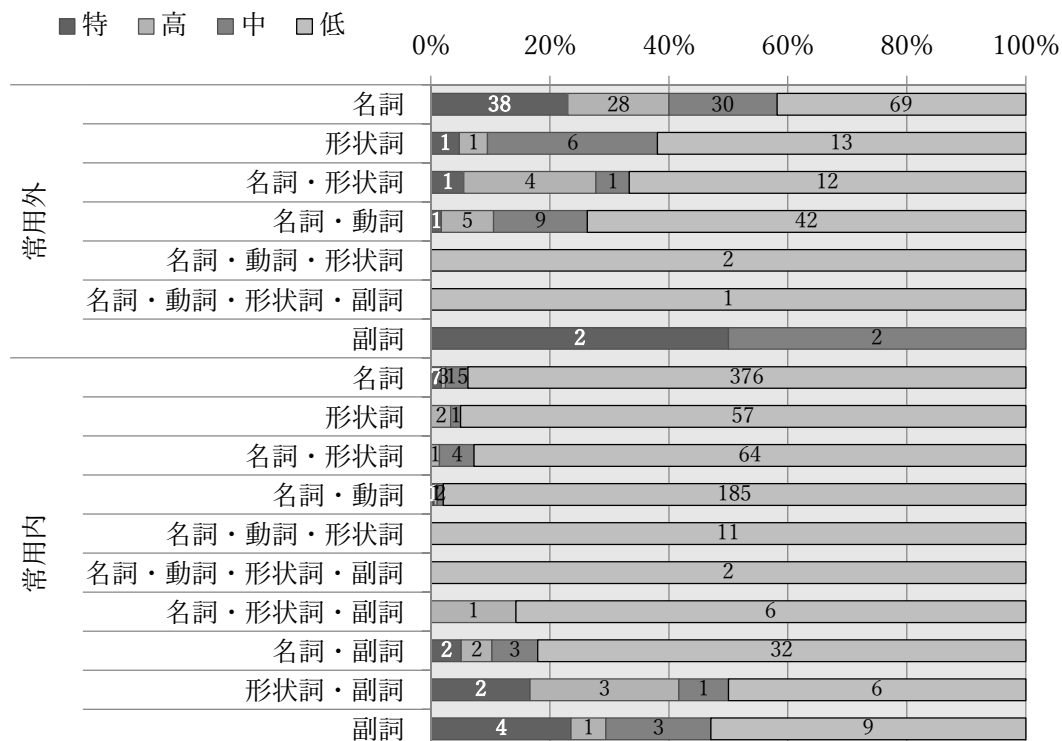


図 6 品詞と仮名表記率の分布

図 6 より、表外字を含む漢語では、動詞用法を含まない語（「勿論(副詞)」「所為(名詞)」「馬鹿(名詞・形状詞)」など）で仮名表記率の高い語の割合がより高く、常用漢字のみから

なる語では「副詞」の用法を持つ漢語（「沢山(形状詞・副詞)」「多分(名詞・形状詞・副詞)」「是非(副詞)」など）で、仮名表記率の高い語の割合が高いことが分かる。ここから、動詞は仮名表記になりやすく、副詞は仮名表記になりやすい傾向にあることが予想される⁵。

ただし、この品詞ラベルはコーパスの形態論情報に基づくもので、個々の漢語の使用実態や用法に基づくものではないため、用法と仮名表記率との相関を確認するために、いくつかの語を取り上げ、実際の用法に基づいて仮名表記率の分布を見てみよう。ここでは、「是非」（表内字）を取り上げ、例 6 のように、格助詞等に接続する、あるいは、例 7 のように連体修飾を受け、「是と非、良し悪し」の意で用いられているものを「名詞」、例 8 のように連用修飾成分になり「強く願う」意を表すものを「副詞」として、用例の分析を行った。

(6) それが不服とおっしゃるならば、天宮において是非を論じられるがよろしい。

【出典】 BCCWJ サンプル ID : PB19_00498, 井上祐美子(著)『乱紅の琵琶』1950 年

(7) 今回参拝を止めた理由と参拝の是非。

【出典】 BCCWJ サンプル ID : OC05_01310, Yahoo!知恵袋, 2005 年

(8) 一度はぜひお目に懸かなければなるまいと思います

【出典】 BCCWJ サンプル ID : PB32_00008 杉山秀子(著)『プロメテウス』2003 年

■漢字（是非） ■平仮名（ぜひ） ■その他（片仮名） ■漢字（早速） ■平仮名（さっそく） ■その他

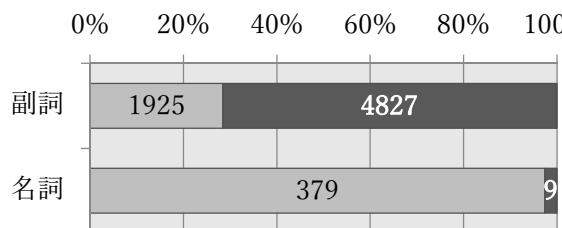


図 7 「是非」の用法と表記の相関

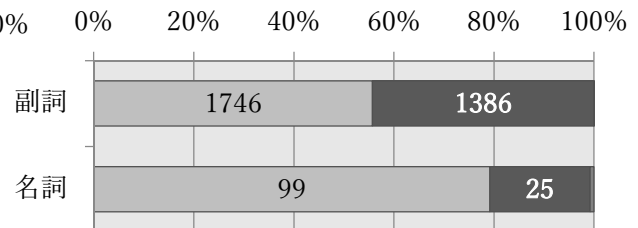


図 8 「早速」の用法と表記の相関

図 7 より、常用漢字のみからなる漢語「是非」は、名詞用法では 97%で漢字表記が選択されているのに対して、副詞では漢字表記 28%に対して、仮名表記が 71%を占め表記傾向が逆になっている。また、「是非」のように用法による意味の相違が明確なもの以外の例として、「早速」（常用内、語彙レベル B、仮名表記率中）を取り上げ、「早速やる」のような副詞用法と、「早速ですが」「早速に」「早速の回答」のような名詞用法における表記を確認すると、図 8 に示す通り、同様に副詞用法でより仮名表記が選択される傾向にあることが分かる。

このように、副詞において仮名表記が選択される傾向は、例えば漢字使用に関する指針となりうる「公文における漢字使用及び送り仮名の付け方について」（昭和 57 年 1 月 14

⁵ 漢語動詞は、構成要素となる漢字の字義によって想起される意味と語義が緊密であるため漢字表記を捨てにくい、漢語副詞の多くは、元の語義からの意味変化により副詞用法が生まれており、構成漢字の字義と語義の結びつきが希薄であること、また、漢語動詞は実質的な意味を持つものに対して、漢語副詞は程度や時間などの意味を付加する機能的な性質が強いため、機能語や形式名詞・補助動詞等と同様、仮名表記が選択されやすいなどの背景があるものと思われる。

日例規（総）第2号）において「原則として、仮名で表記する副詞」として挙げられる「かなり ふと やはり よほど」に含まれない語においても多く見られる傾向であり、漢字か仮名かという表記の選択が、①常用漢字か否かという知識ベースの区分に寄らず、②用いられ方（品詞）差により行われている点は、3.5 節に指摘した、意味分野による統一的な表記傾向と同様、統一性・合理性を志向した表記法の一つと意味づけることができる。

3.7 レジスターと仮名表記率の相関

前節まで、BCCWJ の全体に対する表記傾向の分析を行ったが、採取した用例がどのような媒体に掲載されたものかによって、表記の傾向が異なる可能性があるため、本節では、コーパスの付加情報「レジスター」を用い、使用媒体と表記の関係について検討する。

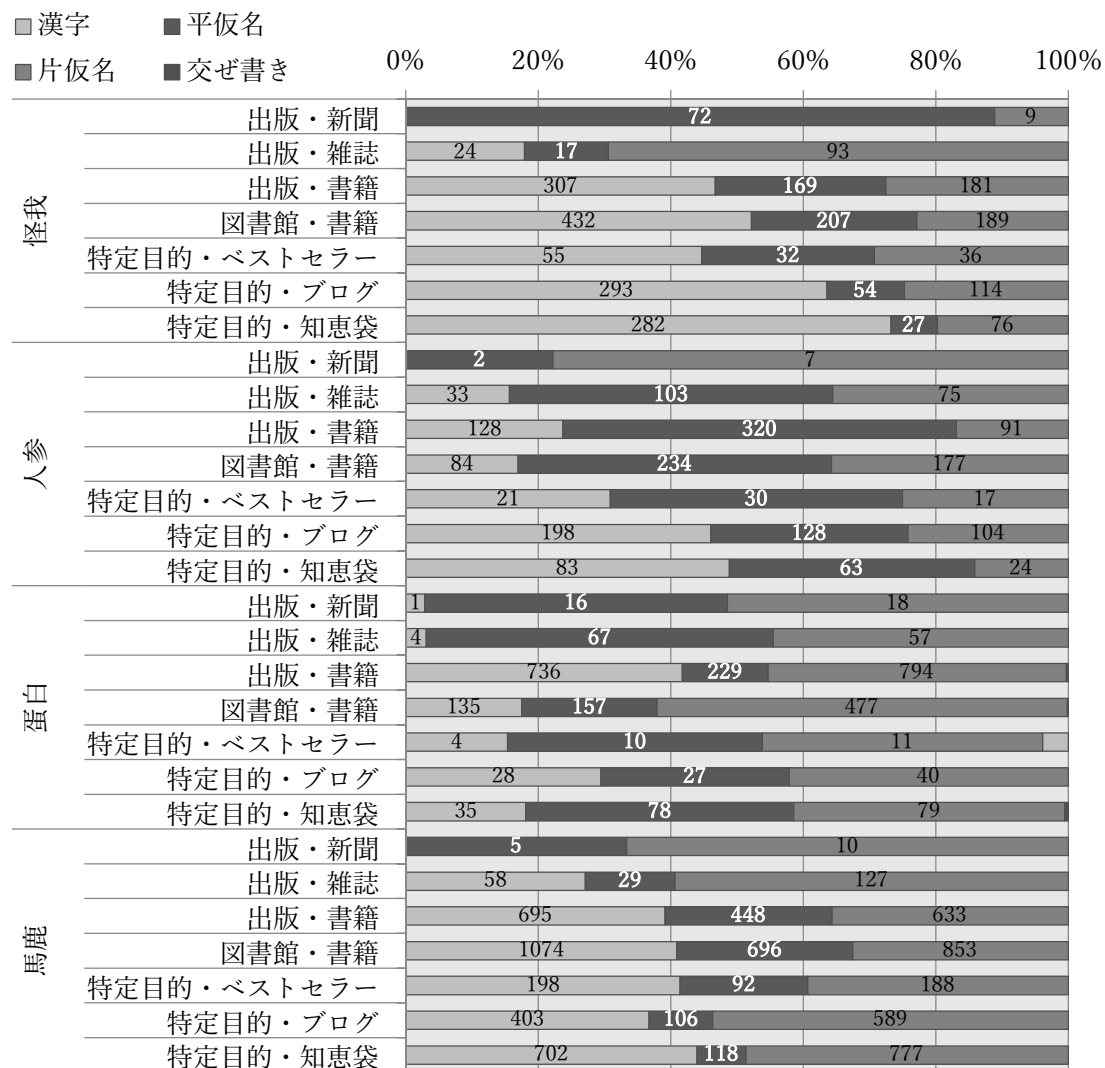


図9 表外字を含む漢語のレジスター別表記分布

調査対象とした 1075 語から、適切な用例数が確保できる語彙レベル B（語彙レベル A は表記の揺れが少なく、また用例数が多すぎるため分析が困難であるため除く）の語で、仮名表記率が「高」の語から、表外字（音）を含む「馬鹿」「蛋白」「怪我」「人參」、常用漢

字のみからなる「多分」「是非」「随分」「大抵」を取り上げ、表記実態を調査した⁶。

表外字を含む語について、図9を見ると、語によるばらつきがあるものの、概観すれば、新聞ではほぼ100%に近く仮名表記が用いられ、次いで雑誌、書籍と漢字表記率が増えWeb媒体のブログ・知恵袋では他より漢字表記が多いという傾向が見て取れる。一方、常用漢字のみの語について図10を見ると、語によって新聞、雑誌、書籍の分布は様々だが、総じてWeb媒体のブログと知恵袋で漢字表記が多い。

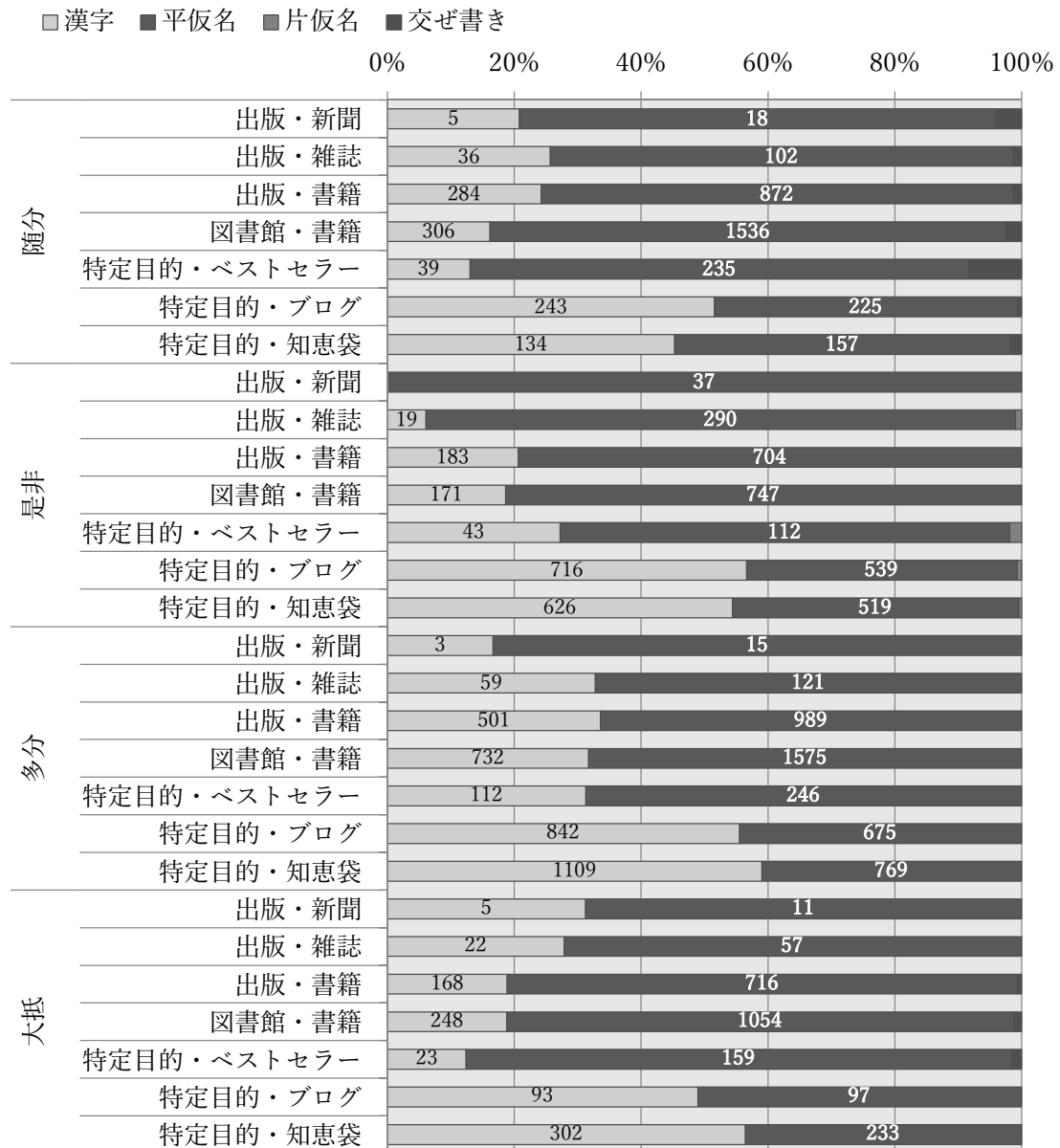


図10 常用漢字のみからなる漢語のレジスター別表記分布

また、図11として、先に品詞による表記傾向の異なりを示した「是非」について、レジ

⁶ 特定目的SCのうち国会会議録、広報誌、韻文、法律、白書は、調査対象語を含まない（あるいは極めて低頻度）場合があるため、分析対象から除外した。

スターの差を含めた表記分布を示す。品詞による表記の使い分けについても、新聞・雑誌ではこれが顕著であり、書籍類でも明確な差が見られる。一方 Web 媒体のブログ・知恵袋では、用法の差は見られるものの、副詞用法における仮名表記は半数以下であり、他の媒体に見られる仮名表記の優位性が見られない。

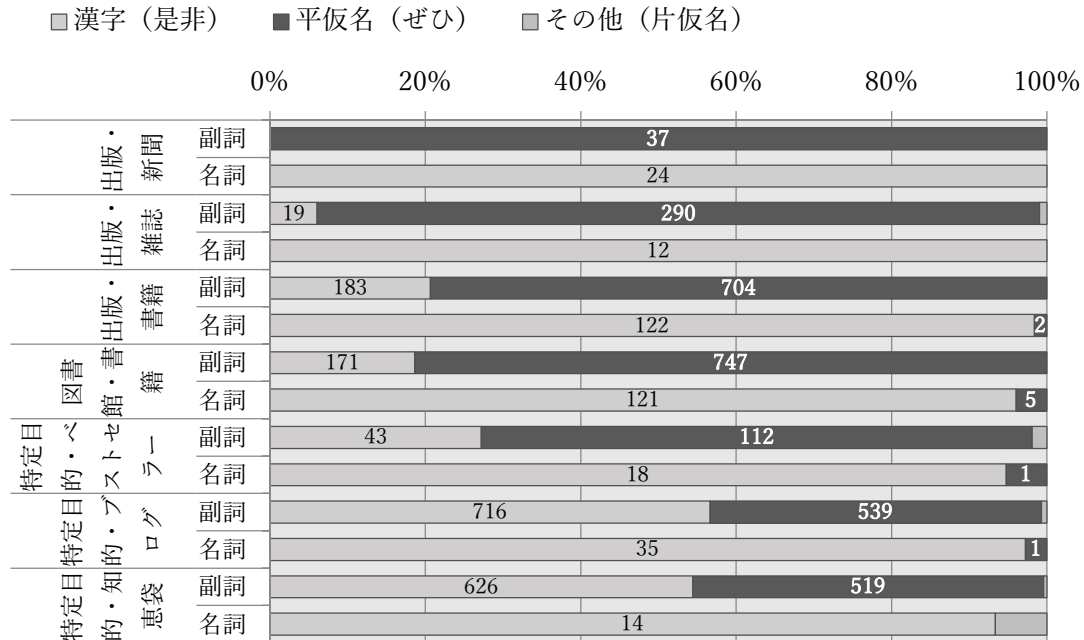


図 11 レジスター別にみた「是非」の品詞別表記分布

ここから、新聞社や出版社の記者・編集者による用字は統制が利いており、表外字や副詞用法を持つ語で平仮名表記の比率が高いが、個々の著者の用字がより反映されやすい書籍においては、表外字の漢字表記が新聞・雑誌より多く、更に標準的表記への志向性や表記統制の影響を受けにくい Web 媒体の用字においては、字体や用法による表記選択の傾向が他の媒体より低いことが分かる。Web 媒体の個人的な用字においては、仮名漢字変換の影響として、変換候補に現れる漢字表記を無意識・無意図的に選択している可能性が高い。

4. まとめ

以上、BCCWJ の網羅的な表記実態調査に基づく漢語の仮名表記率（3.1 節）から漢語の層別を行い、複数の指標との関連性について検討することで、本来漢字で表記される漢語が仮名で表記される背景として、以下の六つの要因があることを指摘した。

- ① 字体特徴：常用漢字表外字（音）を含む語は仮名表記率が高いが、表内字でも仮名表記率の高い語がある。…3.2 節
- ② 語彙レベル：出現頻度の高い漢語ほど仮名表記率は低い。…3.3 節
- ③ 音声変位形の有無：「面倒／メンドウ」に対する「メンド」のような音声転訛形を持つ語は仮名表記率が高い。…3.4 節
- ④ 意味分野：動植物やそれと近接する食物の分野では仮名表記率が高い。…3.5 節
- ⑤ 品詞：副詞用法を持つ語は仮名表記率が高く、複数の品詞があるものは、表記の使い分けがある場合が多い。…3.6 節

- ⑥ レジスター：新聞や雑誌は標準的な表記選択が行われているが、Web 媒体のテキストではそれによらず、仮名表記率が低い傾向がある。…3.7 節

また、仮名表記選択に最も強い影響を与えると思われる①の字体特徴にかかわらず、意味分野や品詞において特定の語彙群が、同様の傾向（字体特徴から予測される標準的な表記から逸脱した表記を選択する）を見せることがあることを指摘し、これらの表記の選択や嗜好には、類似性に基づく合理化作用が働いている可能性を示した（3.5 節，3.6 節）。

5. おわりに

本発表では、漢語の仮名表記の背景について複数の要因を示したが、要因相互の関係性については検討が至らなかった。多変量解析手法などを取り入れ、漢語の仮名表記の要因について更に整理を進めたい。また、発表者は、近代から現代への漢語の変化を研究の主題としており、漢語の仮名表記についても、実態調査から近代と現代とで差異が見られることが分かっている。これについては、機を改めて報告したい。

謝 辞

本稿は、日本学術振興会特別研究員奨励費 16J08872「コーパスを利用した近現代漢語の表記・語法の多様性に関する計量的・通時的研究」（代表：間淵洋子）による成果の一部である。

文 献

- 天野成昭・近藤公久(2000)『日本語の語彙特性—朝日新聞の語彙・文字頻度調査〈第7巻〉頻度 (NTT データベースシリーズ)』三省堂。
- 石井久雄(2001)「ひらがなの文法性・語彙性」『同志社大学留学生別科紀要』1, pp.3-16
- 岩原昭彦・八田武志(2004)「文字言語における感情的意味情報の伝達メカニズムについて」*Cognitive Studies*,11:3, 271-281.
- 柏野和佳子(2014)「『コーパス』でさぐる和語や漢語のカタカナ表記の実態」高田智和・横山詔一編『日本語文字・表記の難しさとおもしろさ』彩流社, pp. 86-105.
- 児玉徳美(2013)「日本語の用字用語」『立命館文学』63, pp.410-392.
- 増地ひとみ(2013)「テレビ番組の文字情報における文字種の選択—番組のジャンルと語用論的要素に注目して—」『早稲田日本語研究』22, pp.24-35.
- 中山恵利子(1998)「非外来語のカタカナ表記」『日本語教育』(日本語教育学会)96, pp. 61-72.
- 成田徹男・榊原浩之(2004)「現代日本語の表記体系と表記戦略—カタカナの使い方の変化—」『人間文化研究』(名古屋市立大学)2, pp. 41-55.
- 則松智子・堀尾佳代子(2006)「若者雑誌における常用漢字のカタカナ表記化—意味分析の観点から—」『北九州市立大学文学部紀』72, pp.19-32
- 寺田博視・田中久美子(2008)「単語親密度と単語頻度の関係に関する一考察」『言語処理学会第14回年次大会発表論文集』, pp.713-716.
- 臼木智子(2008)「雑誌の片仮名表記—基準から外れる表記について—」国学院大学大学院紀要. 文学研究科 40, pp.265-280.

関連 URL

コーパス検索アプリケーション『中納言』

<https://chunagon.ninjal.ac.jp/>